DOCUMENT RESUME

ED 214 394 FL 012 847

AUTHOR Poplack, Shana
TITLE "Sometimes I'll Start a Sentence in Spanish Y TERMINO EN ESPANOL": Toward a Typology of Code-Switching. CENTRO Working Papers, No. 4.
INSTITUTION City Univ. of New York, N.Y. Centro de Estudios Puertorriguenos.
PUB DATE Mar 79
NOTE 83p.; Prepared by Language Policy Task Force.

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS Adults; *Bilingualism; *Code Switching (Language); English; Ethnic Groups; Language Research; Language Usage; Linguistic Competence; *Puerto Ricans; *Sociolinguistics; Spanish; *Spanish Speaking; *Speech Communication; Syntax
IDENTIFIERS New York (New York)

ABSTRACT

This paper attempts to integrate the results of the ethnographic and attitudinal components of a broader study into a specifically sociolinguistic analysis. While a variety of opinions can be found in the literature on code-switching, the contention here is that code-switching is a norm in specific speech situations that exist in stable bilingual communities. Satisfaction of this norm requires more linguistic competence in two languages than has heretofore been noted. The code-switching behavior of 20 Puerto Ricans from 102nd Street in New York City, who had varying degrees of reported and observed bilingual ability, was observed and described. The quantitative analyses are based on recorded speech data in both interview and natural settings. The analysis deals with linguistic questions concerning the surface configuration of the switch, and with the code-switchers themselves. A finding considered to be crucial is that there are virtually no ungrammatical conbinations of the two languages in the 1,835 switches studied, regardless of the bilingual ability of the speaker. It is also shown that switch types traditionally considered most deviant, those switches occurring within a single sentence, are the ones requiring most skill. They tend to be produced by the "true" bilinguals in the sample. (AMH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED214394

CENTRO

WORKING PAPERS

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Nelida Perez
Center for Puerto Rican Studies

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

FL012847

Centro de Estudios Puertorriqueños

"Sometimes I'll Start a Sentence in Spanish Y TERMINO EN ESPAÑOL": Toward a Typology of Code-Switching

4

Language Policy Task Force

New York, March 1979

Copyright ©, RESEARCH FOUNDATION OF
THE CITY UNIVERSITY OF NEW YORK, 1979

APRENDER A LUCHAR
LUCHAR ES APRENDER

CENTRO DE ESTUDIOS PUERTORRIQUENOS 445 W 59th St.; New York, N.Y. 10019 Tel.: (212) 489-5260

"Sometimes I'll Start a Sentence in Spanish Y TERMINO EN ESPAÑOL": Toward a Typology of Code-Switching[1]

Shana Poplack
Language Policy Task Force

1. Introduction

An overwhelming majority[2] of Puerto Ricans residing in the continental United States currently claim Spanish as their 'mother tongue'. This is true for young as well as older speakers, despite the fact that most of them were either born, raised, or spent a good part of their adult life in an English-speaking society.

Along with signs of vigor and renewal of the language, however, there is also some indication that use of Spanish is on the wane, especially among the younger generations of speakers who were born and raised in New York City (Pedraza 1978). The present investigation is part of an interdisciplinary study which aims to examine the place of both Spanish and English in a Puerto Rican community in East Harlem through 1) participant observation of their distribution in the daily life of the community, 2) analysis of attitudes of community members toward each of the languages, and 3) quantitative sociolinguistic analysis of selected linguistic behavior.

Due, among other things, to a circulatory pattern of migration, this appears to be a stable bilingual community, rather than a transitional one where acquisition of a second language would eventually displace the first (Fishman 1971). This pattern of displacement of the mother tongue has characterized several early twentieth century immigrant groups in the United States, and has usually been brought to completion by the third generation. In contrast, the Puerto Rican community under investigation includes third-generation speakers of both Spanish and English.

102nd Street, a block in the heart of El Barrio, perhaps the oldest continuous Puerto Rican settlement in the United States, provides a unique setting to investigate these issues. Block residents are predominantly (95%) Puerto Rican, to the virtual exclusion of all other ethnic groups, an attribute which is not characteristic of the Puerto Rican community in the United States as a whole (Pedraza 1978). In the Spanish language and Puerto Rican

culture are to survive in the United States, their chances of doing so are presumably greatest in such an ethnically homogeneous environment.

This paper is an attempt to integrate the results of the ethnographic and attitudinal components of the broader study into a specifically sociolinguistic analysis.

2. Code-Switching

Long-term ethnographic observation of 102nd Street, carried out by Pedro Pedraza (1978), indicated three modes of communication among block members: English-speaking, Spanish-speaking and code-switching. While use of Spanish predominates in certain domains (such as in the home or while playing numbers), its exclusive use in any of these settings was not observed. Similarly, while use of English predominates in official settings, it is also possible to hear Spanish in these domains. Pedraza further observed that 'there were speakers who code-switched because they lacked full command of Spanish and those who code-switched because they lacked full command of English' (p. 33). However, as we will see, it is only by linking ethnographic observations with linguistic analysis that code-switching behavior may be most adequately explained.

Code-switching is the alternation of two languages within a single discourse, sentence or constituent. In a report on an earlier study of a balanced bilingual speaker (Poplack 1978), code-switching was categorized according to the degree of integration of items from one language (L_1) to the phonological, morphological and syntactic patterns of the other language (L_2).

Because the balanced bilingual has the option of integrating his utterance into the patterns of the other language or preserving its original shape, items such as those in (1) below, which preserve English phonological patterns, were considered examples of code-switching in that study, while segments such as those in (2) which are adapted to Puerto Rican Spanish patterns, were considered to be instances of monolingual Spanish discourse.[3]

1a. Leo un MAGAZINE [mægə'ziyn] (I read a magazine.)

1b. Me iban a LAY OFF [lέy ɔ̀hf] (They were going to lay me off.)

2a. Leo un magazine [maɣa'siŋ] (I read a magazine.)

2b. Me iban a dar layoff ['leiof] (They were going to lay me off.)

In the ensuing sections we explore code-switching on a community-wide basis, focusing on speakers of varying

bilingual abilities. Inclusion in the sample of non-fluent bilinguals requires modifying our previous definition of code-switching. In the speech of non-fluent bilinguals segments may remain unintegrated into L_2 on one or more linguistic levels, due to transference of patterns from L_1. This combination of features leads to what is commonly known as a 'foreign accent', and is detectable even in the monolingual L_2 speech of the speaker, as in (3) below, which was rendered wholly in Puerto Rican Spanish phonology:

3. That's what he said. [da 'waɾi se] (58/100)[4]

In order to consider an utterance such as (3), which occurs in an otherwise entirely Spanish context, as a code-switch from Spanish into English, we have refined the criteria for identifying a code-switch in terms of the type of integration as in Table 1.

Type	Levels of Integration Into Base Language			CS?	Example
	phon	morph	syn		
1	✓	✓	✓	No	Es posible que te MOGUEEN. (They might mug you.) (002/1)
2	-	-	✓	Yes	Las palabras HEAVY-DUTY, bien grandes, se me han olvidado. (I've forgotten the real big, heavy-duty words.) (40/485)
3	✓	-	-	Yes	[da 'wari se] (58/100)
4	-	-	-	Yes	No creo que son FIFTY-DOLLAR SUEDE ONES. (I don't think they're fifty-dollar suede ones.) (05/271)

Table 1. Identification of code-switching according to type of integration into the base language.[5]

The example of Type 1, mogueen, is phonologically, morphologically and syntactically integrated into the base language, although etymologically a loan word from English mug. It is here considered an instance of monolingual Spanish discourse. In contrast, Type 4 segments are totally unintegrated into the patterns of the base language. This sort of code-switch occurs most typically in the speech of balanced bilinguals. Type 2 follows English phonological and morphological patterns, but violates English syntactic patterns. The example shown follows the Spanish syntactic pattern of adjective placement. This type of segment is also considered a code-switch into English, although one which violates the 'equivalence constraint'. (See p. 10 below.) Although Type 3 involves phonological integration into Spanish (i.e. follows Puerto Rican Spanish phonological rules), it is morphologically, syntactically and lexically English. Thus the example of Type 3 is considered a code-switch into English, rendered with a 'foreign accent'. Spontaneous switches of words, sentences and larger units at a turn boundary, not involving any change in interlocutors, were also considered to be code-switches if they exhibited Types 2, 3, or 4 of integration in Table 1.

3. Theoretical Background

Much of the literature on code-switching (e.g. Gumperz 1971, 1976; Gumperz and Hernandez-Chavez 1970, Valdés-Fallis 1976, 1978; McClure 1977) has focused on its social and pragmatic functions. While there is little doubt that functional factors are the strongest constraints on the occurrence of code-switching, it is clear that linguistic factors also play a role. This paper demonstrates how the incorporation of both functional and linguistic factors into a single model is necessary to account for code-switching behavior.

Although in some of the earlier literature (e.g. Lance 1975) the occurrence of code-switching was characterized as random, most investigators now appear to agree that in many aspects it is rule-governed, despite the fact that there is little agreement on the precise nature of the rules involved. Proposed grammatical rules have generally taken the form of categorical constraints based on acceptability judgments of invented instances of code-switching (Gingràs 1974, Timm 1975, Gumperz 1976). While acceptability judgments provide a manageable way to tap community grammar norms, their use is questionable in the case of an overtly stigmatized sociolinguistic marker, as

is the case of code-switching (Gumperz 1971). Moreover, studies of code-switching performance in two widely separated bilingual communities have independently yielded counterexamples to these categorical constraints (Pfaff 1975, 1976· Poplack 1978).

More importantly, the proposed constraints are not of the general nature one would wish to ascribe to linguistic universals.[6] In Poplack (1978) two syntactic constraints on code-switching were suggested, which together were general enough to account for all instances of code-switching in the Puerto Rican data on which that study was based as well as the Chicano data on which the majority of the code-switching literature is based, and at the same time restrictive enough not to generate instances of non-occurring code-switches:

a) The free morpheme constraint. Codes may be switched after any constituent in discourse provided that constituent is not a bound morpheme.[7] This constraint holds true for all linguistic levels but the phonological, for reasons explained above. Thus, a segment such as (4) may be produced, where the first syllable follows the Caribbean Spanish tendency to aspirate /s/

before voiceless consonants, while the second syllable follows English phonological patterns. This should be seen as aiming for, but missing, an English target, rather than a switch between two bound morphemes. However, items such as (5) where the Spanish bound morpheme -iendo (-ing) is affixed to the English root eat, have not been attested in this or any other study of code-switching to my knowledge, unless one of the morphemes has been integrated phonologically into the language of the other.

4. una buena exCUSE [eh'kjuws] (a good excuse)
5. * EAT - iendo (eating)

Included under this constraint are idiomatic expressions, such as cross my fingers (sic) and hope to die and si Dios quiere y la virgen (God and the virgin willing) which are considered to behave like bound morphemes in that they show a strong tendency to be uttered monolingually.

b) The equivalence constraint. Code-switches will tend to occur at points in discourse where juxtaposition of L_1 and L_2 elements does not violate

a syntactic rule of either language, i.e. at points around which the surface structures of the two languages map onto each other. According to this simple constraint, a switch is inhibited from occurring within a constituent generated by a rule from one language which is not shared by the other.[8] This can be seen in Figure 1, where the dotted lines indicate permissible switch points and the arrows indicate ways in which constituents from two languages map onto each other. The speaker's actual utterance is reproduced in (C).

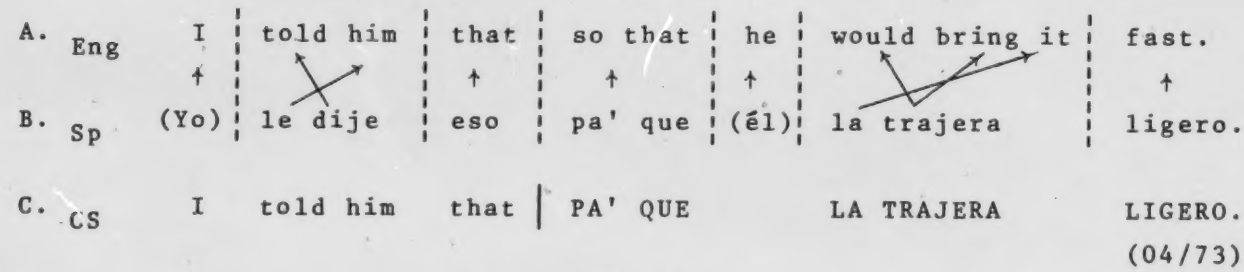

A. Eng I told him that so that he would bring it fast.
B. Sp (Yo) le dije eso pa' que (él) la trajera ligero.
C. CS I told him that PA' QUE LA TRAJERA LIGERO.
(04/73)

Figure 1. Permissible code-switching points.

An analysis based on the equivalence constraint may be applied to the by now classical examples in (6) which were constructed by Gingràs (1974)

and tested for acceptability on a group of Chicano bilinguals.

6a. El MAN que CAME ayer WANTS JOHN comprar A CAR nuevo. (The man who came yesterday wants John to buy a new car.)

6b. Tell Larry QUE SE CALLE LA BOCA. (Tell Larry to shut his mouth.)

Gingràs claims that (6a), in which codes are switched after almost every other word, is 'in some very basic sense different' from (6b), where the switching occurs between major constituents (p. 170). While it is true, as we will see below, that major constituents are switched more frequently than smaller ones, we suggest that constituent size only partially explains the difference between the two sentences, and that the important distinction is with respect to the equivalence constraint.[9] The sentence structures in (6a) and (6b) are similar in that both include a verb phrase and a verb phrase complement. In each, the verbs belong to a class which in English requires that an infinitive complementizer rule apply to the verb phrase complement, while Spanish

makes use of a subjunctive complementizer in this same construction.

Sentence (6a) violates the equivalence constraint because it applies an English infinitive complementizer rule, which is not shared by Spanish, to the verb phrase complement. Because the code-switch did take place in this invented example, an English rule was lexicalized in Spanish, yielding a construction which could not have been generated by a Spanish rule, and which is therefore ungrammatical by Spanish standards. On the other hand, the first portion of (6a) was generated by rules which are shared by English and Spanish (i.e. marked for both L_1 and L_2). The L_1 and L_2 versions map onto each other, constituent-by-constituent and element-by-element. A code-switch may therefore occur at any point within the main clause and the utterance remains grammatical by both L_1 and L_2 standards. Structures in discourse to which a rule from L_1 but not from L_2 must categorically apply were found to be avoided as switch points by the balanced bilingual. There were no cases like (6a) in our

data, and furthermore, all 26 of Gingràs' Chicano informants found it to be unacceptable as well. Constituents whose structures are non-equivalent in L_1 and L_2 tend to be uttered monolingually in actual performance. This occurred in (C) in Figure 1 (as well as in Gingràs' example 6b), where the verb phrase complements have undergone the Spanish subjunctive complementizer rule and are also lexicalized completely in Spanish. Ninety-four percent of Gingràs' informants also found (6b) to be an acceptable utterance.

An additional site of non-equivalence in (6a) is the object noun phrase in the embedded sentence: a car nuevo. English and Spanish have non-equivalent rules for adjective movement. In English, attributive adjectives typically precede the head noun, whereas in Spanish they typically follow it. A closed set of Spanish adjectives may also precede the noun. Switching an adjective but not the noun within the noun phrase, by following either L_1 or L_2 adjective placement rules for adjectives other than those

> in the closed set, results in a construction which is judged unacceptable by Timm's Chicano informants (p. 479) and 'fairly unacceptable' by Gingràs' informants (p. 172). Such constructions occur very rarely in our own performance data.

Simultaneous operation of the free morpheme and the equivalence constraints permits only code-switched utterances which, when translated into either language, are grammatical by both L_1 and L_2 standards, and indicate a large degree of competence in both languages.

One might ask then whether the formulation of these constraints was not simply a consequence of having studied a balanced bilingual speaker: it is not surprising, although it required empirical proof, that an individual with an extensive repertoire in more than one language can manipulate them without violating the grammatical rules of either. But what happens in the case of non-fluent bilinguals? Being clearly dominant in one of the two languages, are they forced to switch into it from time to time because of lack of lexical or syntactic availability when speaking the other?

Weinreich (1953) characterized the ideal bilingual as

an individual who 'switches from one language to the other according to appropriate changes in the speech situation (interlocutor, topics, etc.) but not in an unchanged speech situation, and certainly not within a single sentence'(p. 73, emphasis added). He further speculated that there must be considerable individual differences between those who have control over their switching and those who have difficulty in maintaining or switching codes as required (p. 73).

The phenomenon of code-switching has been a point of contention in assessing community identity. While intellectuals have seen language mixture to constitute evidence of the disintegration of the Puerto Rican Spanish language and culture (e.g. Varo 1971, de Granda 1968), community members themselves appear to consider various bilingual behaviors to be defining features of their identity (Attinasi 1979). The opinion that code-switching represents a deviation from some bilingual 'norm' is also wide-spread in educational circles today (LaFontaine 1975). It is our contention here that code-switching is itself a norm in specific speech situations which exist in stable bilingual communities. Furthermore, as we will demonstrate, satisfaction of this norm requires

considerably more linguistic competence in two languages than has heretofore been noted.

The present study addresses these issues by analyzing the code-switching behavior of twenty Puerto Rican speakers of varying degrees of reported and observed bilingual ability.

4. Hypothesis

As documented in Poplack (1978), a single individual may demonstrate more than one configuration or type of code-switching. One type involves a high proportion of intra-sentential switching, as in (7) below.

7a. Why make Carol SENTARSE ATRAS PA' QUE (sit in the back so) everybody has to move PA' QUE SE SALGA (for her to get out)? (04/439)

7b. He was sitting down EN LA CAMA, MIRANDONOS PELEANDO, Y (in bed, watching us fighting and) really, I don't remember SI EL NOS SEPARO (if he separated us) or whatever, you know. (43/412)

We refer to this as a more complex or 'intimate' type, since a code-switched segment, and those around it, must conform to the underlying syntactic rules of two languages which bridge constituents and link them together grammatically.

Another, less intimate type, is characterized by relatively more tag switches and single noun switches. These are often heavily loaded in ethnic content and would be placed low on a scale of translatability, as in (8b).

8a. Vendía arroz (He sold rice) 'N SHIT. (07/79)

8b. Salían en sus carros y en sus (They would go out in their cars and in their) SNOWMOBILES. (08/192)

Many investigators do not consider switches like those in (8) to represent true instances of code-switching, but rather to constitute an emblematic part of the speaker's monolingual style (Gumperz 1971, Wentz 1977). It will also be noted that their insertion in discourse has few, if any, ramifications for the remainder of the sentence. Tags are freely moveable constituents which may be inserted almost anywhere in the sentence without fear of violating any grammatical rule. The ease with which single nouns may be switched is attested to by the fact that of all grammatical categories, they have been found to be the most frequently switched (e.g. Timm 1975, Wentz 1977).

It was found that the choice of intimate versus

emblematic code-switching is heavily dependent on the ethnic group membership of the interlocutor in the case of the balanced bilingual, who has the linguistic ability to make such a choice. In-group membership favors intra-sentential code-switching, while non-group membership favors emblematic switching. In other words, that type of switch which all investigators agree to be 'true' instances of code-switching was mainly reserved for communication with another in-group member. In considering whether this pattern holds true more generally, we first note that the type of code-switching used by non-fluent bilinguals must be further constrained by bilingual ability. The research described in the following sections examines the extent to which the bilingual competence evinced by the skilled code-switching behavior of a balanced bilingual is shared by the non-fluent bilinguals in the same speech community. This will have ramifications for the possible use of code-switching as an indicator of bilingual ability.

With the inclusion of non-fluent bilinguals in the sample, three alternative outcomes may be hypothesized. The speaker may engage in both intimate (intra-sentential) and emblematic switching, regardless of her competence in

the two languages, thereby running the risk of rendering utterances which will be ungrammatical for L_1, L_2 or both (and hence providing a principled basis for the claim that code-switching represents a deviation from some norm). On the other hand, the speaker may avoid those intra-sentential switches which are syntactically risky. This might assure grammatical utterances. Such results would weaken the claims that code-switching occurs due to lack of availability in L_2. Logically, there is a third possibility, that the speaker will not switch at all. This need not be considered; although two members of the sample switched only once each in 2 1/2 hours of speech, there was no one who did not switch at all.

5. The Sample

5.1 Block Description

102nd Street is located in the heart of El Barrio, one of the oldest and until recently the largest Puerto Rican community in the United States. Although Puerto Ricans have now dispersed to other areas of Manhattan and other boroughs in New York City, we estimate that the population on 102nd Street is still at least 95% Puerto Rican.

The block, which is identified by the census to be one of the lowest socio-economic areas of the City, is largely residential. It consists of 16 three- to five-story tenements housing some 600 people, and is bounded by major avenues on the East and West. There is a small number of commercial establishments on the block, including two bodegas, an alcapurria stand, two Hispanic social clubs, a numbers parlor, a pet store, a vegetable market, and a plumbing supply shop.

Block life is active. It is not uncommon to see groups of people congregating on the stoops, and in the summer months, children play in the streets and adults set up tables on the sidewalks for domino games.

The block has had a stable Puerto Rican population since the 1930's, and today includes third and even fourth generation family members. Although uncharacteristically homogeneous with regard to ethnicity, block residents are heterogeneous with regard to personal history. About half of the block residents were born and raised in Puerto Rico. These are mostly older people, who tend to be Spanish dominant or monolingual. The younger people were generally born and raised in El Barrio, and most are either English dominant or bilingual. This fact and the

low median age of the population (three-fourths are under 45 years of age) match the general demographic characteristics of the New York City Puerto Rican population in general (United States Department of Labor 1975:44).

Extended participant observation of 102nd Street indicates that block residents may be divided into nine social networks, which though not mutually exclusive, are seen to reappear consistently in the public life of the block. The informants for this study are drawn from two of the more closely observed of these networks.

Half of the sample belongs to a network centered around the Gavilanes social club. Participants in this network are linked through friendship or familial ties as well as participation in common club activities. The group is male-dominated, and its members are on public display more than other block residents since they are out on the street for large periods of the day. Members range in age from the early twenties to the fifties and most of the males were employed during the period of observation (1975 - 1977). (Only one female in the present sample had paid employment during this time.) The group includes both the Island-born and raised, who are Spanish dominant, and New York City-born and raised,

who are English-dominant or bilingual. Group members generally accommodate to the older, Island-born members by speaking Spanish.

An additional eight informants belong to a network whose members congregate around the numbers parlor (Banca), a center of lively social activity on the block. Calling the numbers and passing them on by word-of-mouth (in Spanish) is a daily event in block life. Like members of the Gavilanes group, Banca group members range in age from the twenties to the fifties; however, both men and women participate in this network. Friendship and familial ties are less of a factor in linking group members than shared participation in activities in and around the Banca. Members of this network were all born and raised in Puerto Rico, and are Spanish-dominant or bilingual. Banca members were for the most part unemployed during the period of observation.

Of the two remaining informants, one does not participate in any network significantly more than in any other, and the second is a recent arrival to the block. Both are Spanish-dominant.

5.2 The Informants

The twenty informants in the sample were selected

primarily on the basis of two parameters: age of arrival in the United States, and (ethnographically) observed language preference. This choice was made in order to study the effects of bilingual ability and community influence on code-switching. Sample members are fairly evenly divided between those who arrived as children or were born in the United States (0 - 6 years of age), when parental influence on child language use is greatest; those who arrived as pre-adolescents (7 - 12), when peer influence encroaches on parental influence; those who arrived as adolescents (13 - 17), when peer influence on language choice is greatest (Payne 1976, Poplack 1978a); and those who arrived as adults, when patterns of language use tend to have crystallized.

About half of the speakers were observed to be Spanish-dominant; about 10% English-dominant; and the rest bilingual. With some exceptions these observations confirm the speakers' own self-reports. (Cf. Section 5.4 below.)

In addition, several other demographic, ethnographic and attitudinal factors which could affect language use were studied.

5.3 Demographic Characteristics of the Sample

Eleven of the twenty informants are male, nine are female, and 75% of them are between 21 and 40 years of age. Regardless of age of arrival, all but two have spent 10 years or more in the United States, and the majority (60%) has been there for more than 20 years. Since duration of stay is probably overshadowed by other influences on linguistic behavior after about two years (Heidelberger Forschungsprojekt "Pidgin-Deutsch" 1977), the level of competence in English of this group is not likely to change much with increased residence in the United States.

Most of the sample (65%) have also spent 10 or more years in Puerto Rico, either early in life (only one informant was born in the United States), or including extended sojourns after having migrated to the continental United States. There is some correlation between speaker age and time spent in Puerto Rico: the older speakers are generally those who have spent most time on the Island. However, the majority of the sample (70%) visits the Island infrequently: less than every two years. In sum, it is reasonable to expect that these speakers have all also acquired some competence in Spanish, whether it is put to use at present or not.

Sample members report more years of schooling than the general Puerto Rican population in New York City. Seventy percent have had at least some high school education; this includes 20% who have either graduated high school or had some college education.

Slightly less than half the sample was employed at the time of the study, and these mainly in the service sectors of the work force: baker, cook, medical technician, counterman, etc. Of these, the majority (70%) is employed off the block.

5.4 Language-Oriented Characteristics of the Sample

The informants responded to a language attitude questionnaire, designed to tap various aspects of language skill by self-report as well as community attitudes towards Puerto Rican language and ethnicity.

When asked if they considered themselves 'mainly Spanish speakers, English speakers, or bilinguals,' slightly more than half (55%) of the sample claimed to be mainly Spanish speakers. The others considered themselves bilingual. No one claimed to be mainly an English speaker, reflecting the underrating on the part of speakers of their English language skills. There is doubtless some correlation between this finding and the fact that all speakers

report having learned Spanish in early childhood (between ages of 2 and 7), with the majority (75%) having learned it in Puerto Rico. On the other hand, only two informants learned English in early childhood. Sixty percent of the sample learned English between the ages of 7 and 21, and all speakers report having learned English in the United States. When asked to rate their competence in both Spanish and English on a seven-point scale, speakers' ratings are consistent with their verbal self-description. Almost all (95%) rate themselves as having more than median proficiency in Spanish, while less than half (45%) claim this for English. If the hypothesis that code-switching is caused by lack of availability in L_2 were correct, it would appear that members of this sample population would favor switching into Spanish from an English base.

5.5 Attitudinal Characteristics of the Sample

Our language attitude questionnaire also seeks to tap speakers' feelings towards language as it relates to Puerto Rican culture and ethnicity. Responses to attitude questions are less readily interpretable than reports of language use. On an ethnic identity scale based on questions about pride in being Puerto Rican, feelings towards assimilation and speakers' characterization of

nationality (e.g. Puerto Rican, Nuyorican, American, etc.), the majority (65%) revealed clear positive identification as Puerto Ricans. When asked to assess how important the Spanish language is to 'Puerto Ricanness', 90% felt that Spanish is 'important' or 'very important' to being a Puerto Rican. This attitude was summarized by Sally:

> 'SI TU ERES PUERTORRIQUEÑO (if you're Puerto Rican), your father's a Puerto Rican, you should at least DE VEZ EN CUANDO (sometimes), you know, HABLAR ESPAÑOL (speak Spanish).' (34/25)

Nevertheless, a majority (60%) also felt that Puerto Rican monolingual English speakers did not represent a divisive force in the Puerto Rican community in New York City.

When asked if there was 'anything you could say in Spanish that you could not say in English,' or vice-versa, 60% of the speakers felt that there was nothing that could be said in one language that could not be translated into the other. For the most part (75%), the speakers in this sample are also aware that code-switching is a frequent and wide-spread phenomenon in their community. When asked if they thought that 'few, some or many speakers mixed languages,' three-fourths of the sample thought that many

people code-switched. Such awareness of community and individual behavioral norms with regard to code-switching was again voiced by Sally, who provided the title for this paper.

> 'Sometimes I'll start a sentence in Spanish Y TERMINO EN ESPAÑOL [sic] (and finish in Spanish).' (34/489)[10]

This opinion corroborates Pedraza's ethnographic observation that a majority of block residents code-switched somewhat or frequently.

In sum, although the Spanish language is overwhelmingly considered an integral part of being Puerto Rican, and in spite of the fact that only a minority feels that one or the other of the languages may be better for saying certain things, ethnographic observation, quantitative sociolinguistic analysis, and speakers' own self-reports all indicate that code-switching is an integral part of community speech norms on 102nd Street. Code-switches provoked by lack of availability or utilized as an emblem of ethnic identity appear, then, to be only weak factors in speakers' perception of their own behavior.

6. Methodology

6.1 Data Collection

The quantitative analyses which follow are based on recorded speech data in both interview and 'natural' settings. Pedraza's membership in the Puerto Rican community and his familiarity with the setting and participants allowed him to enter local network situations, such as domino games or bochinche (gossip) sessions, and simply turn on his tape-recorder without causing an apparent break in the conversational flow. In addition, he carried out a 'sociolinguistic interview'[11] with each informant aiming to elicit casual, undirected speech, and administered the detailed attitude questionnaire mentioned above. Our data, then, range from the vernacular speech of informal, intra-group communication, to the more formal discourse used in discussing concepts such as language and culture.

The importance of data collection techniques cannot be overemphasized, particularly in the study of a phenomenon such as code-switching, which cannot be directly elicited. The actual occurrence of a switch is constrained, probably more than by any other factor, by the norms or the perceived norms of the speech situation. The most important of these norms for the balanced bilingual

was found (Poplack 1978) to be the ethnicity of the interlocutor, once other criteria (appropriateness, formality of speech situation) were met. The balanced bilingual speaker switched four times as frequently with an in-group interlocutor as with a non-member and, what is more, used a much larger percentage of intimate switches with the in-group member. Since the data utilized in this study are representative of in-group interaction only, they are presumably characterized not only by a higher rate of code-switching, but also by a larger proportion of intra-sentential switches, than would have been the case if they had been collected by a non-member.

6.2 Coding Procedures

Sixty-six hours of tape-recorded speech in which each informant participated in each of the three speech situations provided 1,835 instances of code-switching. Transcription, coding and analysis of the data were carried out with the invaluable collaboration of Alicia Pousada. The informants appeared on tape, either alone or in a group, from a minimum of 2 1/2 hours to a maximum of 11 hours. Each instance of a switch was coded as to its syntactic function in the utterance, along with the syntactic categories of the segments which preceded

and followed it. We also noted the language of the switch, whether it was preceded or followed by editing phenomena (hesitations, false starts, etc.), whether it constituted a repetition of the preceding syntactic category, whether it was a single noun switch in an otherwise L_2 base, and if so, whether it was an ethnically loaded item.

In each instance, the largest complete constituent next to the switch in the syntactic derivation of the utterance was considered to be the syntactic category of the segment adjacent to the switch. Thus, con los puños in (9a) was coded as a prepositional phrase preceded by an independent clause and followed by a tag. Promising in (9b) was coded as a verb preceded by an auxiliary and followed by an object noun phrase, since all of them are dominated by the same verb phrase node. Que in (9c) was coded as a subordinate conjunction preceded by an independent clause and followed by a subordinate clause. This method was used to determine syntactic category of the code-switch and segment following it as well.

9a. But I wanted to fight her CON LOS PUÑOS (with my fists), you know. (43/356)

9b. Siempre está PROMISING cosas. (He's always promising things.) (04/408)

9c. I could understand QUE (that) you don't know how to speak Spanish, ¿VERDAD? (right?). (34/24)

Other segments of varying syntactic make-up, but which exercised a consistent function in discourse, were coded according to that function. Examples of these are fillers, e.g. este (umm), I mean; interjections, e.g. ¡ay, Dios mio! (oh, my God!), shit!; tags, e.g. ¿entiendes? (understand?), you know; idiomatic expressions, e.g. y toda esa mierda (and all that shit), no way; quotations, e.g. put down 'menos' (less). These are segments which are less intimately linked with the remainder of the utterance, insofar as they may occur freely at any point in the sentence. As will be seen, they contrast with the intra-sentential switches, which must obey sentence-internal syntactic constraints.

Certain switched segments were larger than a single constituent, as in (10) below. Hasselmo (1970) has called this 'unlimited switching'. These cases, which were relatively rare, required special coding conventions. The only type with non-negligible frequency involved moveable

constituents like sometimes and honey in (10). Since these constituents do not form an integral part of the syntactic structure of the sentence, they were relegated to the category of intervening material between the switch and the adjacent syntactic categories, and the switches were considered to have occurred between the independent clause and the adverb in (10a), and between the verb phrase and the adverbial phrase in (10b).

10a. No tienen ni tiempo (they don't even have time) SOMETIMES FOR THEIR OWN KIDS, AND YOU KNOW WHO I'M TALKING ABOUT. (04/17)

10b. Se sentó (he sat down), HONEY, AWAY FROM US. (04/433)

In a case like (11), utterances were divided by sentence boundary. Pa' muchos sitios was coded as a switched prepositional phrase in Spanish preceded by an independent clause and followed by a sentence. With my husband was coded as a switched prepositional phrase in English preceded by an independent clause and followed by a pause.

11. And from there I went to live PA' MUCHOS SITIOS (in a lot of places). Después viví en la ciento diecisiete (then I lived on 117th) WITH MY HUSBAND. (42/76)

It will be noted that the analysis in (11) involves

a change of base language. The first prepositional phrase is considered a switch into Spanish from an English base, while the second is considered a switch into English from a Spanish base. While speakers who are dominant in one language show a strong tendency to switch into L_2 from an L_1 base, more balanced bilinguals often alternate base languages within the same discourse.[13] An example of this can be seen in (12) which represents a single discourse, and where segments to the left of the slashes exhibit a base language different from those to the right of the slashes.

11. But I used to eat the BOFE, the brain. And then they stopped selling it because TENIAN, ESTE, LE ENCONTRARON QUE TENIA (they had, uh, they found out that it had) worms. I used to make some BOFE! / Después yo hacía uno d'esos (then I would make one of those) CONCOCTIONS: THE GARLIC con cebolla, y hacía un mojo, y yo dejaba que se curara eso (with onion, and I'd make a sauce, and I'd let that sit) FOR A COUPLE OF HOURS. / Then you be drinking and eating that shit. Wooh! It's like eating anchovies when you're drinking. Delicious! (04/101)

6.3 Non-Code-Switches

Certain borderline alternations between L_1 and L_2 were excluded from this study. One type involves switched items which have been referred to by Hasselmo (1970) as socially integrated into the language of the community:

segments which are repeated often enough in a certain language to be regarded as habitualized. These may or may not be phonologically integrated into the base language; and should not be confused with the types of integration shown in Table 1:

13a. Ay, ¡qué CUTE [kju] se ve! (How cute he looks!) (34/202)

13b. Eso es un TEAM [tiɳ]: 'Palo Viejo'. (That's a a team: Palo Viejo.) (37/42)

13c. En ese tiempo había muchos JUNKIES [jɔŋki]. (At that time there were a lot of junkies.) (34/40)

Switches into L_2 designating food names, proper names and place names were also omitted from this study, except when there was an acceptable L_2 alternative which was not used, e.g. Puerto Rico [pɔɤə ˊɤiykow]∿[pwɛrtɔ ˊxikɔ].

Also excluded were translations in response to requests for information, as in (14a), L_2 segments followed by an explanation in L_1, as in (14b), switches accompanied by metalinguistic comments, as in (14c), and instances of 'externally conditioned switching' (Clyne 1972:70) in which the interlocutor switched languages within the same discourse and the informant followed suit, as in (14d).

14a. A: Lo pusieron un...¿cómo se dice? ¿un tutone? (They gave him a--how do you call it? tuton?)

B: TUTOR? (52/229)

14b. But I used to eat the BOFE, the brain. (04/101)

14c. I'm one of those real what you call in Spanish PENDEJAS (jerks), you know. (04/158)

14d. A: I had a dream yesterday, last night.

B: ¿DE QUE NUMERO? (What number?)

A: EL CERO SETENTA Y CINCO. (Zero seventy-five.) (34/040)

6.4 Quantitative Analysis of the Data

We are concerned here with both linguistic and extra-linguistic questions, and we shall attempt to incorporate the answers into a single analytic model. The linguistic questions concern the surface configuration of the switches. Are there some sorts of constituents in discourse which can be switched and others which cannot? Are there constituents which tend to be switched into one language rather than the other? In what ways do switched items combine with unswitched portions of discourse?

The extra-linguistic questions concern the code-switchers. Can the community as a whole be characterized by some code-switching type, or are there speakers who favor certain switch types over others? In the latter

case, what are the demographic, attitudinal, and social factors which contribute to the occurrence of one type over another, and what is the comparative effect of each?

To answer these questions the following quantitative analyses were performed on the data. The syntactic category of the switched item was crosstabulated first with the preceding and then the following syntactic category to ascertain whether certain points (as for example, the point between determiner and noun) in discourse were more favorable to the occurrence of a switch than others. Also crosstabulated were switched item by language, to see if certain switch types were favored by one language over another, and switched item by speaker, to see if there was any difference in switching behavior among speakers.

The crosstabulations revealed that speakers could be divided into two groups: one which favored extra-sentential switches, and another which tended towards the intra-sentential, or more intimate type. The code-switching data were subsequently collapsed into two categories: intra-sentential and extra-sentential switches. These categories were then crosstabulated individually with the demographic, attitudinal and language-

oriented characteristics of the informants to discover which, if any, have an incidence on the choice of one code-switch over the other. Tests of association were applied to each of these crosstabulations to determine the significance of the extralinguistic factors for the occurrence of one code-switch type over the other.

Having thus been able to determine a dependent variable (in this case, code-switch type), its relevant variants (intra-sentential and extra-sentential code-switching), and the total population of utterances in which the variation occurs (i.e. the total corpus of switched items), the significance tests made it possible to suggest which extralinguistic factors might reasonably be expected to affect the relative frequency of the two types of code-switches. Because these factors may be correlated among themselves within the sample and have correlated effects on code-switching type, it was then necessary to use multivariate statistical techniques to determine which factors made a significant contribution, independent of the effects of other factors, to the choice of code-switch type. Because of the binomial nature of the data and their uneven distribution among the different possible configurations of factors, a

maximum likelihood approach was taken for the evaluation of factor effects, together with log-likelihood tests of significance.

7. Results

Perhaps the most striking result of this study is that there were virtually no instances of ungrammatical combinations of L_1 and L_2 in the 1,835 switches studied, regardless of the bilingual ability of the speaker.

Our hypotheses as to the nature of syntactic constraints on code-switching in the speech of the balanced bilingual, i.e. the free morpheme constraint and the equivalence constraint, are generally corroborated by the present investigation of both balanced and non-fluent bilinguals. There were no examples of switches between bound morphemes of the type *eat -iendo mentioned in (5) above. A small number (five, or less than 1% of the data) of switches within idiomatic expressions did occur, however, as in (15a), where a Spanish expression is broken up, and (15b), where an English expression is broken up.

15a. Estamos como marido y WOMAN. (We are like man and wife. < Sp. Estamos como marido y mujer.) (05/141)

15b. Mi mai tuvo que ir a firmar y SHIT pa' sacarme, YOU KNOW. (My mom had to go sign 'n shit to get me out, you know.) (07/058)

A small number of switches which violated the equivalence constraint also occurred (11, or less than 1% of the data). The majority (7/11) of these involved adjective placement, a rule which is not shared by L_1 and L_2. This can be seen in (16a), where adjective placement follows Spanish but not English rules, and (16b), where the reverse is true.

16a. Tenían patas flacas, pechos FLAT. (They had skinny legs, flat chests.) (09/432)

16b. I got a lotta BLANQUITO (whitey) friends. (34/274)

A strong tendency to avoid non-equivalence is nonetheless manifested in the fact that 88% (49/56) of the adjectival forms in the corpus are either predicate adjectives, which have equivalent surface structures in Spanish and English, or members of the subset of Spanish adjectives which precede the noun, as in English.

The proportion of switched items which when combined with the rest of the utterance did not follow grammatical

rules shared by both L_1 and L_2 is negligible. (Of this small number, none of the constructions was idiosyncratic, or based on rules which were not drawn from one or the other of the grammars.) This finding is strong evidence that alternation between two languages requires a high level of bilingual competence. Code-switching involves enough knowledge of two (or more) grammatical systems to allow the speaker to draw from each system only those rules which the other shares, when alternating one language with another. Surprisingly enough, this knowledge appears to be shared by even the non-fluent bilinguals in the sample. The way in which these latter speakers are able to fulfill the requirement of grammaticality, despite their limited competence in one of the codes, will be examined in section (7.5) below.

7.1 Discourse Functions of Code-Switching

The finding that code-switching constitutes the skilled manipulation of overlapping sections of two (or more) grammars is further corroborated by an examination of some of the ways in which a code-switch functions in discourse. One of the characteristics of skilled code-switching is a smooth transition between L_1 and L_2 elements, unmarked by false starts, hesitations or lengthy

pauses. When we examine the data we find that the transition between the preceding category and the switched item is made smoothly, i.e. with no editing phenomena, 96% of the time, while the transition between switched item and following syntactic category is made smoothly 98% of the time. Other characteristics of skilled code-switching include a seeming 'unawareness' of the alternation between languages, i.e. the switched item is not accompanied by metalinguistic commentary, it does not constitute a repetition of all or part of the preceding segment, nor is it repeated by the following segment; switches are made up of larger segments than just single nouns inserted into an otherwise L_2 sentence; and code-switching is used for purposes other than that of conveying untranslatable items.

On examining the data for the sample as a whole, we find that these characteristics are strongly in evidence: the switched item only constitutes a repetition of the preceding segment 5% of the time, while the following segment repeats all or part of the switched item only 8% of the time. Single noun switches constitute only 10% of the date; of these, less than one-fourth represent items which are ethnically loaded.

In other words, features known to be characteristic

of communication with a non-group member, such as high percentages of single noun switches used to convey notions which are difficult to translate, are not defining features of intra-group communication.

7.2 Linguistic Properties of Switched Segments

Having established that switching occurs in a smooth fashion, we turn our attention to the nature of the switches themselves. Which constituents are switched, and in what ways do they combine with preceding and following segments? Do certain combinations tend to occur more regularly?

Fifteen syntactic categories whose occurrence is dependent on sentence-internal constraints were extracted from the data, along with seven extra-sentential, or freely distributable categories. These appear in Table 2 below.

	Syntactic Category of CS	# of CS from Eng. to Sp.	# of CS from Sp. to Eng.	% of Total CS	N
INTRA-SENTENTIAL	determiner	3	0	0.2%	3
	(single) noun	34	141	9.5	175
	subj. noun phrase	44	25	3.8	69
	obj. noun phrase	62	78	7.6	140
	auxiliary	0	0	0.0	0
	verb	6	13	1.0	19
	verb phrase	27	13	2.2	40
	indep. clause	44	35	4.3	79
	subordinate (& relative) clause	53	23	4.1	76
	adjective	3	12	0.8	15
	pred. adjective	6	37	2.3	43
	adverb	14	33	2.6	47
	preposition	2	0	0.1	2
	phrases (prep., adj., advb., inf.)	55	39	5.1	94
	conjunctions (sub-ordinate, coordinate, relative pronoun)	33	16	2.7	49
EXTRA-SENTENTIAL	sentence	201	171	20.3	372
	filler	9	11	1.1	20
	interjection	26	89	6.3	115
	idiomatic expression	8	23	1.7	31
	quotation	20	14	1.9	34
	tag	9	403	22.5	412
	TOTALS	659	1176		1835

Table 2. Code-switching by syntactic category and language.[14]

The relative frequencies with which constituents may be switched, indicated in Table 2, largely confirm the findings of other studies (Gumperz 1976, Wentz 1977, Poplack 1978). As can be seen, full sentences are the most frequently switched constituent, making up 20% of the data. Extra-sentential code-switching types, which require less knowledge of two grammars since they are freely distributable within discourse, together constitute about half the data.

Among the intra-sentential switches, we find single nouns to be the most frequently switched category, again confirming the findings of other studies. Table 2 also reveals a tendency to switch major constituents, which account for about 60% of the intra-sentential data, more frequently than smaller ones. This provides additional support for the equivalence constraint, which predicts that whole constituents will be switched rather than elements within them if the syntactic rule for generating the constituent is not shared by both L_1 and L_2.

7.3 Language of the Switch

Table 2 also shows the frequencies with which the syntactic categories under investigation are produced in each language. As can be seen, with a few exceptions,

segments are about as likely to be switched into English as into Spanish, providing further evidence for the suggestion that the code-switching mode proceeds from a single grammar.

Examining the data in Table 2 more closely, we may test whether the rate of occurrence of a given syntactic category of a switch is significantly different from one language to the other. For this we compared a log-likelihood of a rate estimate for the two languages separately compared to that for the combined data. Significantly more switches from Spanish into English were found for four categories: tags, interjections, single nouns, and predicate adjectives. The latter is probably an artifact of sparse data, but the results for tags, interjections, and single nouns have important interpretations.

It is not surprising that bilinguals residing in an English-speaking society should favor English noun switches over Spanish. Interjections and tags, as will be shown in section (7.5.), are precisely the switch types which are favored by Spanish-dominant speakers. These speakers not only switch almost uniquely into English from a Spanish base, but are also distinguishable from the bilinguals by the type of constituents they switch.

The statistical analysis shows that aside from the four switch types which are favored from Spanish into English, most of the rest tend to be switched significantly more from English into Spanish. This result is an artifact of the other, however, and when the tag, interjection, and noun switches are removed from the data, almost all of the remaining switches show no significant rate differences between the two languages.

7.4 Combinability of Switched Segments

In order to ascertain points within the sentence at which segments may be switched, intra-sententially switched items were crosstabulated with segments preceding and following them. In a 225-cell table generated by the 15 possibilities of syntactic category preceding the code-switch versus 15 syntactic categories for the code-switch itself, about 40% were filled; i.e. 88 different combinations of some constituent and another switched one occurred. Of the non-occurrent combinations, 40% were syntactically impossible (e.g. auxiliary + preposition); the remainder of the cells were empty due most probably to the distribution of the data--cells corresponding to relatively rare switch types and/or relatively rare preceding categories would not be expected to occur with this sample size.

The two most frequently recurring switch points among 681 tokens of intra-sentential code-switch + preceding category were between determiner and noun (19%) and between verb phrase and object noun phrase (12%). This is not surprising since we already have evidence that nouns and noun phrases are frequently switched. Other combinations which recur frequently include independent clause and subordinate clause (4%), verb and predicate adjective (4%), and subject noun phrase and verb phrase (3%). The remaining combinations each represent 2% or less of the data. Similarly, 63 (28%) of the possible combinations occurred among the 729 tokens of code-switch + following category, the overwhelming majority of which also individually represented very small proportions of the data.

Because of the size of the data set, it is statistically unlikely that clearer patterns could emerge from such a large-celled table. With more data we might be able to predict frequencies with which code-switched items precede and follow specific constituents. We hypothesize that such frequencies would simply reflect the frequency of any given combination of constituents (e.g. adverb + adjective, preposition + noun phrase) in monolingual

speech. This hypothesis could only be confirmed by evaluating the frequencies of all the possible constituent combinations in a large sample of monolingual speech, a task beyond the scope of the present research.

What information we do have, however, indicates that there is a rather large number of points within the sentence at which it is permissible to switch codes. This is additional evidence that code-switching requires knowledge of two systems. Note that there is about as much intra-sentential as extra-sentential switching (Table 2) in the corpus. While extra-sentential switching could presumably be accomplished by alternately drawing on rules from two separate grammars, intra-sentential code-switching would appear to depend on the juxtaposition of constituents too intimately connected to be generated separately by rules from two distinct grammars. This, together with the finding that only a very small number of switches are accompanied by breaks in the speech flow, lend strong support to the hypothesis that code-switching is in fact a verbal mode distinct from English-speaking and Spanish-speaking, yet which consists of the overlapping elements from both.

7.5 Differential Behavior of Informants: Language of the Switch

Let us now examine the individual code-swtiching behavior of the informants in the sample.

Table 3 shows the frequency with which speakers switch into English. It is striking that the Spanish-dominant speakers switch almost uniquely into English from an unambiguously Spanish base. Bilingual speakers on the other hand, cluster around the half-way mark, with some switching somewhat more into Spanish from an English base, others the reverse. It is clear that these bilinguals cannot be said to have a single base or dominant language of discourse, but rather two.

Note that three bilingual and two Spanish-dominant speakers (according to self-reports) show patterns which contrast with the other members of their respective groups. As will be seen again in section (7.5.1) below, four of these are speakers whose self-report of language dominance conflicted with our ethnographic and linguistic observations in this regard. The fifth, Edo, is in fact bilingual, but has strong feelings towards speaking Spanish, and has been observed to do so almost uniquely when interacting on the block.

SPANISH-DOMINANTS

INFORMANT	% OF CS FROM SP. TO ENG.	N
Eli	100%	9
Gui	100	35
Tera	100	1
Isi	100	45
Rosa	100	1
Fela	97	69
Charlie	94	33
Sami	93	93
Chito	92	89
Shorty	71	40
Wilda	63	127
Average	92%	542

BILINGUALS

INFORMANT	% OF CS FROM SP. TO ENG.	N
Cal	100%	35
Edo	99	212
Apache	91	63
Pearl	57	35
Garra	53	15
Candy	52	81
Lola	43	309
Melo	39	89
Sally	37	354
Average	63%	1293

N = 1,835

Table 3. Percentage of code-switches into English for Spanish-dominant and Bilingual speakers.

7.5.1 Differential Behavior of Informants: Switch Type

Let us now examine another way in which bilinguals differ from Spanish-dominant speakers.

The switches in Table 2 were listed according to the presumed degree of bilingual proficiency required to produce them, in decreasing order. Lowest on the scale are tag-like switches. These include interjections, fillers, tags, and idiomatic expressions, all of which can be produced in L_2 with only minimal knowledge of the grammar of that language. Next on the scale are full sentences or larger segments, which require much more knowledge of L_2 to produce, although hypothetically, not as much as is required by the third category, intrasentential switches. As suggested above, in order to produce this latter sort of switch, the speaker must also know enough about the grammar of each, and the way they interact, to avoid ungrammatical utterances.

Figure 2 graphs the percentages of each of these switch types for the informants in our sample.[15]

FIGURE 2

PERCENTAGES OF SWITCH TYPES FOR SPANISH-DOMINANT & BILINGUAL SPEAKERS

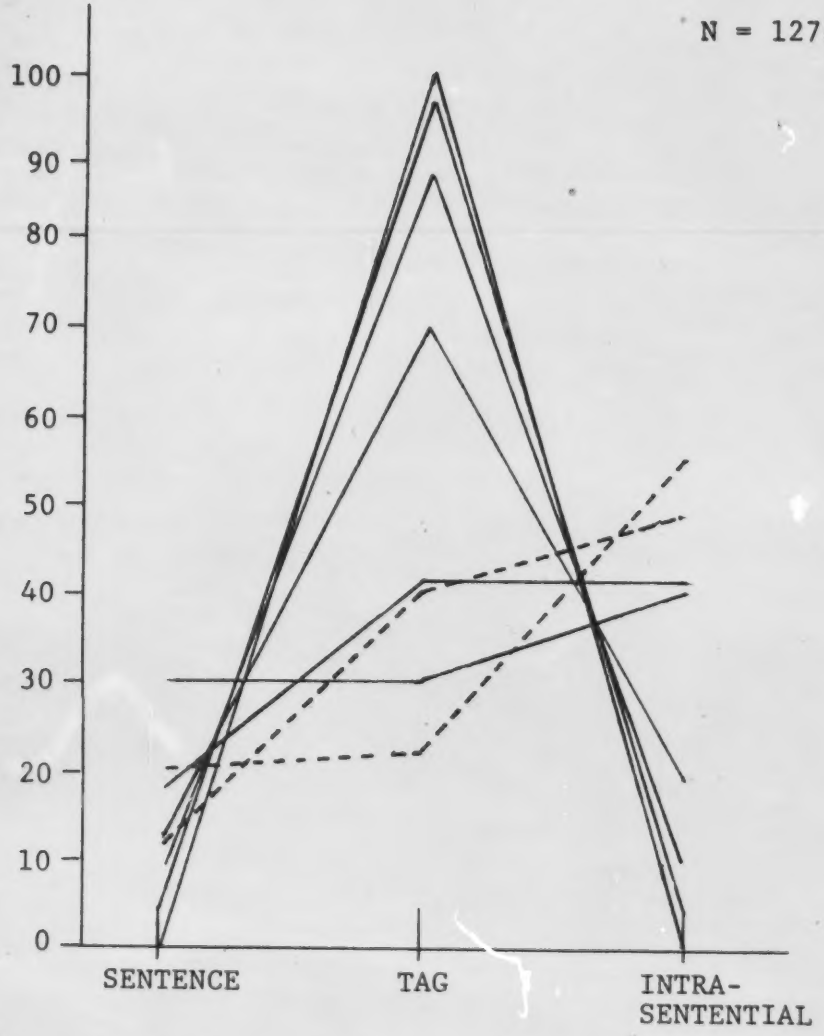


Figure 2a. Percentage of Switch Types for Reported Spanish-dominant Speakers.

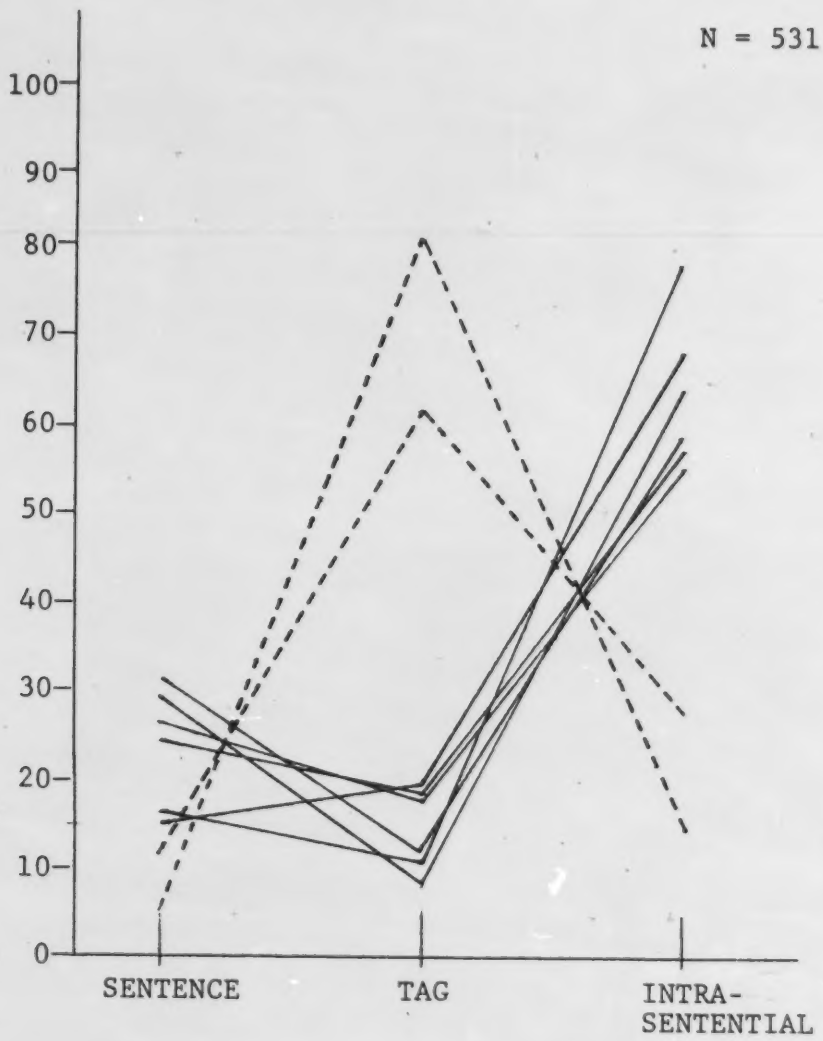


Figure 2b. Percentage of Switch Types for Reported Bilingual Speakers.

Figure 1 shows that reported language ability (which in all cases but four corresponds to observed ability) is an excellent indicator of code-switching patterns. Figure 2a shows that most of those who report that they know, feel more comfortable in, and use more Spanish than English, tend to switch into L_2 by means of tag-like constructions, sometimes to the practical exclusion of sentential or intra-sentential switches. Those who claim to be bilingual, on the other hand, show a reversal (Figure 2b). They favor large amounts of the switches hypothesized to require most knowledge of both languages, sentential and intra-sentential switches. The most favored switch type for bilinguals is clearly intra-sentential, while the least favored is tag-like switching.

The few exceptions to these patterns are represented by the dotted lines on the graphs. Two speakers who claimed to be Spanish-dominant, in fact show a similar code-switching configuration to the bilinguals, while two who claim to be bilingual show patterns similar to the Spanish-dominants. Strikingly enough, these are precisely the cases where ethnographic observation and linguistic analysis were previously found to conflict with self-report, because the speakers underrated or overrated their

ability in English. Two additional speakers are actually Spanish-dominant, but also show a code-switching pattern similar to that of the bilinguals. They do, however, have a greater degree of competence in English than those who follow the 'Spanish-dominant' pattern. Their switches tend to distribute among the three switch types, rather than show a marked preference for any one, and they may be considered to exhibit code-switching behavior intermediate to the bilinguals and the more clearly Spanish-dominant speakers.

8. Contribution of Extralinguistic Factors to the Occurrence of Code-Switch Type

Having ascertained that reported and observed bilingual ability is an important factor in predicting the type of code-switch that will be uttered (a chi-square test shows this factor to be significant at the .001 level), we next attempted to determine which other extralinguistic factors might have an effect on the occurrence of intra-sentential code-switching. Individual tests were performed on each factor group as well as each pair of factors within each group to determine their significance.

8.1 Sex

The sex of the speaker is a significant factor in predicting code-switch type at the .001 level. Women favor intra-sentential switching; over half (56%) of their switches are intra-sentential, while only one-third of the men's switches are of this type.

8.2 Age of L_2 Acquisition and Age of Migration to the United States

These two factors were originally examined separately for each speaker. All informants learned Spanish in early childhood, though there was some variation in the age at which English was learned. Since there is a one-to-one correspondence between the age at which the speaker arrived in the United States and the age at which she learned English (not a surprising fact since all speakers report having learned English in the United States), these two factors were subsequently considered as one. On ethnographic grounds we had originally distinguished four ages of arrival/L_2 acquisition: early childhood (2 - 7), pre-adolescent (8 - 13), adolescent (14 - 18), and adult (over 18). Tests of association reveal that the difference between learning English/arriving in the United States as an adolescent or

as an adult is not significant at the .05 level. Consequently, these two age groups were considered together. All the other age distinctions were significant at either .001 or the .01 levels.

Those speakers who learned both English and Spanish in early childhood, the 'true' bilinguals, show the highest percentage of intra-sentential switching (346/582, or 59%). Those who learned English between the ages of 8 and 13, show only a slightly lower percentage of this type of switch (309/593, or 52%), a small difference, but one which is nonetheless significant at the .01 level. Speakers who learned English after the age of 13, however, show a much lower percentage of intra-sentential switches (196/660, or 30%), a drop which is significant at the .001 level.

8.3 Reported Bilingual Ability

As we have already seen in Figure 1, reported (and observed) bilingual ability is an excellent predictor of code-switching type. Bilinguals produce a far greater percentage of intra-sentential switches (682/1293, or 53%) than those who are Spanish-dominant (169/542, or 31%). This difference is significant at the .001 level.

8.4 Education

Sample members were divided into three categories according to their educational attainment: those who had no more than primary school education (first through eigth grades), those who had some high school but did not graduate (ninth through eleventh grades), and those who graduated high school and/or attended some college. An initial test performed on this factor group revealed that it has an effect on code-switching which is significant at the .001 level. Speakers who had primary school education or less tended to switch intra-sententially slightly more (45%) than those who had some high school education (41%). This difference is not significant. High school graduates, however, switched intra-sententially more (60%) than either of the other two groups, a significant difference at the .001 level. Upon closer examination, however, it became apparent that although the category contained six speakers, all but five of the tokens (401/406) had been uttered by two bilingual women who had learned English as children. Sex, bilingual ability and age of L_2 acquisition are the three factors we have already seen to be highly correlated with choice of switch type.

8.4 Age

Sample members were divided into two groups: younger speakers, those between 20 and 40, and older speakers, those over 41. Speaker age was found not to be significant in predicting code-switch type.

8.5 Social Network Membership

All but two informants belong to two ethnographic networks we called the Gavilanes and the Banca groups, both of which include both Spanish-dominant and bilingual speakers. Network membership is not a significant factor in predicting code-switch type.

8.6 Ethnic Identity

People were divided into those who scored low on a composite index measuring positive feelings towards Puerto Rican identity and those who obtained high scores. Those with positive feelings towards Puerto Rican ethnic identity switched codes intra-sententially somewhat more (446/881, or 51%) than those who had negative feelings (405/954). This difference is significant at the .001 level.

8.7 Continued Contact with Puerto Rico

Sample members were also divided according to relative frequency of return visits to Puerto Rico. Those who return to the Island, where they must speak more Spanish than in New York, more frequently than once every two years, tend to switch intra-sententially far less (71/304, or 23%) than those who visit Puerto Rico less frequently than once every two years (780/1576, or 49%). This difference is significant at the .001 level. Further perusal of these categories, however, indicates that all speakers who migrated to the United States in early childhood (and hence tend towards bilingualism) return to Puerto Rico less frequently than once every two years, and that all those who return more frequently are Spanish-dominant, a distinction we have already seen to be significantly correlated with a low incidence of intra-sentential switching.

8.8 Work Place

A distinction was made between those informants who are employed off the block, where they must presumably interact in English, and those who are either unemployed or employed on the block, where they may communicate in English, Spanish or code-switching. Those who

work off the block switch intra-sententially only about a third of the time (167/460), while those who remain on the block engage in this switch type half of the time (684/1375), a difference which is significant at the .001 level. It is conceivable that those who spend the better part of the day in a speech situation where code-switching is not appropriate get less practice in switching and are therefore less skilled at it. More likely, however, since only Spanish-dominant speakers are employed off the block, it is this characteristic which undoubtedly accounts for the low percentage of intra-sentential switching by these speakers.

9. Multivariate Analysis of Code-Switch Type

Situations in which the data is poorly distributed, as is the case for the factors of education, continued contact with Puerto Rico, and work place, can be as misleading. As we have seen, the apparently significant effect of one factor may really be mainly due to another, language dominance, a fact which is not brought out by looking at one factor at a time. Multivariate analysis can, within limits, separate out these overlapping

effects, and extract the independent contributions proper to each of several related factors (Rousseau and Sankoff 1978, Sankoff and Labov 1978). It can also provide information on the statistical significance of the distinctions defining each factor group.

Let us now examine the comparative contribution of each extra-linguistic factor found by the factor-by-factor analysis to be significant in the prediction of code-switch type: sex, language dominance of speaker, age of arrival in the United States and age of L_2 acquisition, educational attainment, the speakers feelings towards his own ethnicity, amount of continued contact with Puerto Rico, and location of work place.

To what extent are these extra-linguistic factors significant when considered simultaneously? To carry out this multivariate analysis we used VARBRUL 2 (Sankoff 1975) to calculate factor effects and significance levels of factor groups. The factor effects combine to give the probability that a switch will be intra-sentential, according to the model $\frac{p}{1-p} = \frac{p_o}{1-p_o} \times \frac{p_1}{1-p_1} \times \ldots \times \frac{p_n}{1-p_n}$, where p_o is a corrected mean parameter and $p_1, \ldots, p_n$ are the parameters representing the effects of factors 1, ...,n characterizing

a given speaker. Factor probabilities vary between Ø and 1, with figures higher than .5 favoring intra-sentential code-switching, and figures lower than .5 favoring extra-sentential code-switching. The higher the figure, the greater the contribution to rule application, so comparisons can easily be made between various factors and factor groups. The program also calculates the log-likelihood of the model under a given configuration of factor groups. Different analyses can then be compared to see whether the various factor groups contribute significantly to explaining the differential use of code-switching types among speakers.

In the preceding section we found seven factors that, when examined one by one, seemed to significantly affect code-switch type. Some of these factors, however, are clearly correlated among themselves. To see which ones could be considered to have an independent contribution to choice of code-switch type, we carried out 128 separate analyses, each one corresponding to a different combination of explanatory factors. By examining the log-likelihoods first of the seven one-factor analyses, then the 21 two-factor analyses, and so on, we could detect which factors contributed a significant independent effect to the

explanation of the variation in the data.[16] The result of this was the four-factor analysis depicted in Table 4. Each of the four factors here has a strongly significant independent effect on choice of switch type, but the addition of any of the others: ethnic identity, education or continued contact with Puerto Rico, does not significantly increase the explanatory power of the model.

Corrected mean: .36

SEX	AGE OF ARRIVAL/ L_2 ACQUISITION	LANGUAGE DOMINANCE	WORK PLACE
FEMALE .59	CHILD .55	BILING. .68	OFF-BLOCK .67
MALE .41	PRE-AD. .55	SPANISH .32	ON-BLOCK .33
	ADOLESC. .40		

Table 4. Contribution of extra-linguistic factors to the occurrence of intra-sentential code-switching.

The comparative effect of the factors is of particular interest. As can be seen, reported and observed language dominance is the single factor which most affects the occurrence of this switch type. Bilinguals favor it the most, at .68, while those who are Spanish-dominant disfavor it, at .32. An almost equal effect is shown by

the factor of work place. Those who work off the block where they must use English show a greater tendency to switch intra-sententially, at .67, than those who are unemployed or work on the block, where any of the three modes may be used. Table 4 shows that those speakers who acquired both languages in early childhood or pre-adolescence, a defining feature of the balanced bilingual, switch intra-sententially more (.55) than those who learned L_2 at a later age.

Table 4 also indicates that women, who are often in the vanguard of linguistic change, favor intra-sentential switching more than men.

10. Discussion

We have shown how to incorporate both linguistic and extra-linguistic factors into a single analytical model to account for code-switching performance. The linguistic constraints on this phenomenon, the free morpheme and equivalence constraints, represent the basis on which the foregoing analysis was carried out.

The extra-linguistic factors contribute variably to the occurrence of one switch type over another. Choice of these factors arises from long-term familiarity with

the members of the speech community, and on the basis of attitudinal and ethnographic studies carried out in the community.

An elementary, but crucial finding of this study is that there are virtually no ungrammatical combinations of L_1 and L_2 in the 1,835 switches studied, regardless of the bilingual ability of the speaker. This corroborates the hypothesis as to the nature of syntactic constraints on code-switching advanced in Poplack (1978), for both balanced and non-fluent bilinguals.

By showing that non-fluent bilinguals are able to code-switch frequently, yet maintain grammaticality in both L_1 and L_2 by favoring emblematic or tag-switching, we have also demonstrated empirically that code-switching is not monolithic behavior. Three types of code-switching emerge in the speech performance studied, each characterized by switches of different levels of constituents, and each reflecting different degrees of bilingual ability. Multivariate analysis of extra-linguistic factors confirms that those speakers with the greatest degree of bilingual ability ('true' bilinguals) most favor intra-sentential code-switching, the type we had hypothesized to require most skill. The two Chicano code-switching studies of Gingràs (1974) and Pfaff (1975,

1976) also indirectly support these findings. Gingràs tested a group of Chicano and non-Chicano bilinguals on the acceptability of a series of constructed intra-sententially code-switched utterances. The Chicano bilinguals showed much higher rates of acceptance of his grammatical code-switches than the non-Chicanos. This led him to posit that there were probably code-switching norms peculiar to the Chicano community. While speech communities may be characterized by different code-switching norms, it was also the case that the Chicano group had learned both languages in early childhood, while the non-Chicano informants all learned English as adults. Bilinguals of the first type are precisely those we have shown to engage most in intra-sentential code-switching, a fact which concords with their high rate of acceptance of such switches. Similarly, of three samples of Chicano speakers studied, Pfaff found that those who engaged most in 'deep-s' (intra-sentential) switching were those whose speech was characterized by use of both Spanish and English.

Previous work (Gumperz 1971, 1976; Valdés Fallis 1978, Poplack 1978) has shown that code-switching may be used as a discourse strategy to achieve certain interactional

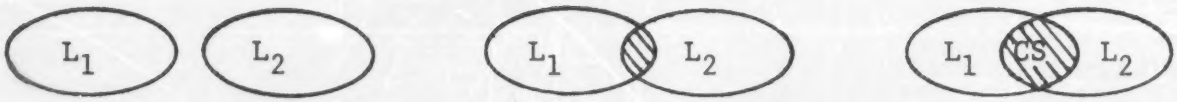
effects at specific points during a conversation. The findings in the present paper, together with the ethnographic observation that code-switching is a linguistic norm in the Puerto Rican community, suggest that this use is characteristic only of certain types of code-switching, which we call 'emblematic', including tags, interjections, idiomatic expressions, and even individual noun switches. On the other hand, a generalized use of intra-sentential code-switching may represent instead an overall discourse mode. The very fact that a speaker makes alternate use of both codes, itself has interactional motivations and implications, beyond any particular effects of specific switches. Indeed, speaker attitudes toward use of Spanish, English and code-switching reported in section (5.4) above do not offer any ready explanation for why a particular segment in discourse should be switched. McClure and Wentz (n.d.) have pointed out that 'there is apparently no real social motivation for, or significance attached to, the practice of code-switching [for example] subject pronouns alone' (p. 266). We agree that there is no 'good reason' (Wentz 1977, e.g. pp. 143, 218) for switching subject pronouns, or lone determiners, etc. Nonetheless, such segments are switched

by bilingual speakers. It may well be possible in some cases for the analyst to impute situational motivations or consequences to specific intra-sentential switches, but the evidence presented here suggests that this has little if any pertinence for the speakers themselves. More important, there is no need to require any social motivation for this type of code-switching, given that, as a discourse mode, it may itself form part of the repertoire of a speech community. It is then the choice (or not) of this mode which is of significance to participants rather than the choice of switch points. When these conditions are met, any segment in discourse may be switched, depending on the bilingual ability of the speaker, and provided it obeys the equivalence constraint. Thus, we cannot agree that switches of, say, object pronouns are 'non-sentences' of any language because 'they violate the social motivation for code-alternation in the first place' (McClure n.d.:265). They simply violate the equivalence constraint!

The suggestion that code-switching is itself a discrete mode of speaking, possibly emanating from a single code-switching grammar composed of the overlapping sectors of the grammars of L_1 and L_2, is supported by several

findings in the present study. We have shown that there is a large number of permissible switch points in the data, rather than a few favored ones. Switching any given constituent in discourse does not necessarily entail continuation in the language of the switch, unless the surface structures are not equivalent in L_1 and L_2. Hence larger constituents are switched more frequently than smaller ones. It was additionally shown that all constituents are about as likely to be switched into L_1 as into L_2, with the few exceptions contingent upon the bilingual ability of the speaker.

In light of these findings, code-switching behavior may be used to measure bilingual ability. Bilingual speakers might have expanding grammars of the type depicted in Figure 3, representing greater degrees of bilingual acquisition. Further empirical studies of code-switching performance in other bilingual communities would provide comparative data to test these hypotheses.



a. Inter-sentential switching b. 'tag'-switching c. Intra-sentential switching

Figure 3. Representation of bilingual code-switching grammars.

These findings taken together and interpreted in terms of the equivalence constraint, provide strong evidence that code-switching is a verbal skill requiring a large degree of linguistic competence in more than one language, rather than a defect arising from insufficient knowledge of one or the other. The rule-governed nature of code-switching is upheld by even the non-fluent bilinguals in the sample. Their behavior suggests at least enough passive competence in L_2 to switch codes by means of the few rules they know to be shared by both languages. It is also striking that precisely those switch types which have traditionally been considered most deviant by

investigators and educators, those which occur within a single sentence, are the ones which require the most skill. They tend to be produced by the 'true' bilinguals in the sample: speakers who learned both languages in early childhood, and who have most ongoing contact with the monolingual English-speaking world.

Code-switching, then, rather than representing deviant behavior, is actually a suggestive indicator of degree of bilingual competence.

NOTES

[1] This analysis is part of a research project on Inter-Generational Perspectives on Bilingualism supported by the National Institute of Education under NIE-G-78-0091, and by the Ford Foundation. The work reported here has benefited from many fruitful discussions with the Language Policy Task Force as well as other researchers at the Center for Puerto Rican Studies. I am particularly indebted to Alicia Pousada for her contribution to all phases of this study. Frank Bonilla, Don Hindle and David Sankoff generously contributed their time and insight, for which I am very grateful. Thanks also to Migdalia Rodriguez for her patience and expertise in the preparation of the text. A portion of this paper was presented at the Linguistic Society of America annual meeting, December 1978.

[2] 92%. (Bureau of the Census 1973)

[3] The reverse pattern, the insertion into an English base of Spanish items with English phonological or morphological patterns is non-existent in this community.

[4] Numbers in parentheses identify speaker and code-switch.

[5] We follow Hasselmo (1970) in designating as the 'base' language that language to which a majority of phonological and morphological features of discourse can be attributed.

[6] For example, why should it be possible to switch codes between a subject and a verb, but not if that subject is pronominal, as suggested by Gumperz (1970) and Timm (1975)?

[7] This constraint is confirmed by data from independent studies (Pfaff 1975, 1976; McClure n.d., Wentz 1977).

[8] A condition similar to the equivalence constraint has been independently suggested by Lipski (1977).

[9] Gingràs also claims that it is not obvious whether (6a) should even be considered an example of code-switching, mainly because 'in the formation of the complement it is not clear whether English transformations have applied in an otherwise Spanish structure' (p. 168). Whether or not an invented example should be considered a code-switch is questionable; in any event, it is one which is hardly likely to occur in actual speech, as operation of the equivalence constraint demonstrates.

[10] This lapsus provides ground for interesting speculation, which we leave to the reader.

[11] This questionnaire was developed for the Philadelphia speech community by the Project on Linguistic Change and Variation under the direction of William Labov, and subsequently adapted for use in the Puerto Rican community by the author.

[12] Note that the sentence itself contains a switch. The switched segment was only coded for following syntactic category if the category was produced in the language other than that of the switch. So an example like Pa' muchos sitios was not coded for following syntactic category.

[13] A construct such as 'language of the sentence', which according to Wentz (1977:182) is the one in which the determiner and main verb were produced, does not appear to be operative for these data, as they contain a good number of code-switched verbs (60, or 3% of the data) in a language other than that of the determiner.

[14] Nouns and verbs were counted as noun phrases or verb phrases respectively if they functioned as such within the utterance.

[15] Three Spanish-dominant and one bilingual speaker were omitted from these calculations as they each produced 15 tokens or less.

[16] I would like to thank David Sankoff for making available a version of the variable rule program which facilitates this stepwise multiple regression procedure.

REFERENCES

Attinasi, John. 1979. Results of a Language Attitude Questionnaire Administered Orally to an Ethnographically Chosen Sample of 91 Residents of a Block in East Harlem, New York, 1977-1978. ms.

Bureau of the Census. 1973. Puerto Ricans in the United States. Publication PC(2)-1E.

Clyne, Michael. 1972. Perspectives on Language Contact. Melbourne:The Hawthorne Press.

Fishman, Joshua. 1971. The Sociology of Language: An Interdisciplinary Approach. in J. Fishman (ed.). Advances in the Sociology of Language. Vol. 1. The Hague:Mouton.

Gingràs, Rosario. 1974. Problems in the Description of Spanish-English Intra-Sentential Code-Switching. in G.A. Bills (ed.). Southwest Areal Linguistics. San Diego:Institute for Cultural Pluralism.

Gumperz, John J. 1971. Bilingualism, Bidialectalism and Classroom Interaction. in Language in Social Groups. Stanford:Stanford University Press.

_____. 1976. The Sociolinguistic Significance of Conversational Code-Switching. Working Papers of the Language Behavior Research Laboratory. #46. Berkeley:University of California.

_____, and Eduardo Hernández-Chávez. 1970. Cognitive Aspects of Bilingual Communication. in E. Hernandez-Chavez et al. (eds.). El Lenguaje de los Chicanos. Arlington:Center for Applied Linguistics.

Granda, Germán de. 1968. Transculturación e interferencia lingüística en el Puerto Rico contemporáneo (1898-1968). Bogotà:Instituto Caro y Cuervo.

Hasselmo, Nils. 1979. Code-Switching and Modes of Speaking. in G. Gilbert (ed.). Texas Studies in Bilingualism. Berlin:Walter de Gruyter and Co.

_____. 1972. Code-Switching as Ordered Selection. in E. Finchow et al. (eds.). Studies for Einar Haugen. The Hague:Mouton.

Heidelberger Forschungsprojekt "Pidgin-Deutsch". 1977. The Acquisition of German Syntax by Foreign Migrant Workers. in D. Sankoff (ed.). Linguistic Variation: Models and Methods. New York:Academic Press.

LaFontaine, Hernán. 1975. Bilingual Education for Puerto Ricans: ¿si o no? Paper presented at the National Conference on the Educational Needs of the Puerto Rican in the United States. Cleveland, Ohio.

Lance, Donald. 1975. Spanish-English Code-Switching. in E. Hernandez-Chavez et al. (eds.). El Lenguaje de los Chicanos. Arlington:Center for Applied Linguistics.

Lipski, John. 1977. Code-Switching and the Problem of Bilingual Competence. in M. Paradis (ed.). The Fourth Lacus Forum. Columbia, S.C.:Hornbeam Press.

McClure, Erica. n.d. The Acquisition of Communicative Competence in a Bicultural Setting. NIE Grant NE-G-00-e-0147 final report.

_____. 1977. Aspects of Code-Switching in the Discourse of Bilingual Mexican-American Children. Washington, D.C.:Georgetown University Roundtable on Languages and Linguistics.

_____, and James Wentz. 1975. Functions of Code-Switching Among Mexican-American Children. in Papers from the Parasession on Functionalism. Chicago: Chicago Linguistics Society.

Payne, Arvilla. 1975. The Acquisition of the Phonological System of a Second Dialect. Unpublished University of Pennsylvania Ph.D. dissertation.

Pedraza, Pedro. 1978. Ethnographic Observations of Language Use in El Barrio. Unpublished ms.

Pfaff, Carol. 1975. Syntactic Constraints on Code-Switching: A Quantitative Study of Spanish/English. Paper presented at the Linguistic Society of America annual meeting.

_____. 1976. Functional and Structural Constraints on Syntactic Variation in Code-Switching. Papers from the Parasession on Diachronic Syntax. Chicago:Chicago Linguistic Society.

Poplack, Shana. 1978a. Dialect Acquisition Among Puerto Rican Bilinguals. Language in Society Vol. 7, No. 1.

_____. 1978. Quantitative Analysis of Constraints on Code-Switching. Center for Puerto Rican Studies Working Paper #2. To appear in R. Durán (ed.) Latino Language and Communicative Behavior. New Jersey:Ablex Publishing Corp.

Rousseau, Pascale, and David Sankoff. 1978. Advances in Variable Rule Methodology. in D. Sankoff (ed.). Linguistic Variation: Models and Methods. New York: Academic Press.

Sankoff, David. 1975. VARBRUL 2. Unpublished program and documentation.

_____, and William Labov. 1978. On the Uses of Variable Rules. To appear in Language in Society.

Timm, L.A. 1975. Spanish-English Code-Switching: el porqué y How-Not-To. Romance Philology. Vol. 28, No. 4.

United States Department of Labor. 1975. A Socio-Economic Profile of Puerto Rican New Yorkers. New York:Bureau of Labor Statistics.

Valdés-Fallis, Guadalupe. 1976. Social Interaction and Code-Switching Patterns: A Case Study of Spanish/English Alternation. in G. Keller et al. (eds.). Bilingualism in the Bicentennial and Beyond. New York: Bilingual Press.

_____. 1978. Code-Switching as a Deliberate Verbal Strategy: A Microanalysis of Direct and Indirect Requests Among Bilingual Chicano Speakers. To appear in R. Durán (ed.). Latino Language and Communicative Behavior. New Jersey: Ablex Publishing Corp.

Varo, Dr. Carlos. 1971. Consideraciones antropológicas y políticas en torno a la enseñanza del 'Spanglish' en Nueva York. Río Piedras:Ediciones Librería Internacional.

Weinreich, Uriel. 1953. Languages in Contact. The Hague: Mouton.

Wentz, James. 1977. Some Considerations in the Development of a Syntactic Description of Code-Switching. Unpublished University of Illinois at Urbana-Champaign Ph.D. dissertation.